Université de Genève

Départment d'informatique

Éléments de la Théorie de l'information

Théorie de l'information pour la science des données et l'apprentissage automatique — Série d'exercices 3

16.03.2022

Délai : 22.03.2022

Note : Veuillez suivre les instructions suivantes
`https://moodle.unige.ch/mod/page/view.php?id=171850`.

Les **Inégalités de Concentration** quantifient la manière dont une variable aléatoire X s'écarte de sa moyenne μ. Elles prennent généralement la forme d'une *borne* pour les queues de X telle que

$$\Pr\{|X-\mu| > t\} \leq \text{something small},$$

où $t > 0$. L'inégalité de concentration la plus simple est l'inégalité de Bienaymé-Tchebychev (que nous avons expliquée dans la série d'exercices précédente).

Le Théorème Central Limite de Lindeberge-Lévy: Soit $X_1, X_2, \ldots$ une séquence de variables aléatoires i.i.d. de moyenne μ et de variance σ^2. Considérer la somme $S_N = X_1 + X_2 + \ldots + X_N$ de N variables aléatoires, et la normaliser pour obtenir une variable alatoire de *moyenne nulle* et de *variance unitaire* come suit:

$$Z_N \triangleq \frac{S_N - \mathbb{E}[S_N]}{\sqrt{\operatorname{Var}(S_N)}} = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^{N} (X_i - m).$$

Alors, pour $N \to \infty$,

$$Z_N \to \mathcal{N}(0,1) \quad \text{en distribution.}$$

La convergence en distribution signifie que la CDF de la somme normalisée Z_N converge point par point vers la CDF de la distribution normale standard $\mathcal{N}(0,1)$. Nous pouvons exprimer cela en termes de queues de distribution comme suit. Pour tout $t \in \mathbb{R}$, pour $N \to \infty$ nou avons

$$\Pr\{Z_N \geq t\} \to \Pr\{G \geq t\} = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-\frac{x^2}{2}} \mathrm{d}x,$$

où $G \sim \mathcal{N}(0,1)$ est une variable aléatoire de loi normale.

Un cas particulier remarquable du théorème central limite est celui où les X_i sont des variables aléatoires de Bernoulli dénotées par $X_i \sim \operatorname{Ber}(p)$ avec un paramètre fixe $p \in (0,1)$. Rappelons que dans ce cas, X_i prend les valeurs 1 et 0 avec respectivement une probabilité p et $1-p$. Rappelons également que $\mathbb{E}[X_i] = p$ et $\operatorname{Var}(X_i) = p(1-p)$. La somme $S_N = X_1 + \ldots + X_N$ est dite avoir la distribution Binomiale dénotée $\operatorname{Binom}(N,p)$. Le théorème central limite permet de dire que pour $N \to \infty$, nous avons

$$\frac{S_N - Np}{\sqrt{Np(1-p)}} \to \mathcal{N}(0,1) \quad \text{en distribution.}$$

Ce cas particulier du théorème central limite est appelé *théorème de Moivre-Laplace.*

Problème 1 (**Un exemple de l'Inégalité de Bienaymé-Tchebychev pour la Distribution Binomiale — 20 points**).

Lancer une pièce de monnaie N fois. Quelle est la probabilité d'obtenir "au moins" $\frac{3}{4}N$ faces ?

Note : Dans ce problème, nous voulons montrer que la probabilité $\Pr\{S_N \geq \frac{3}{4}N\} \leq \frac{k}{N}$. Cela signifie que cette probabilité converge vers zéro *au moins linéairement* avec N.

Indice : En utilisant l'inégalité de Bienaymé-Tchebychev, trouver la quantité k.

Problème 2 (**Queues de la Distribution Normale — 40 points**).

Soit $Z \sim \mathcal{N}(0,1)$. *Prouver* que pour tout $t > 0$,nous avons

$$\left(\frac{1}{t} - \frac{1}{t^3}\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \leq \Pr\{Z \geq t\} \leq \frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

Rappel de l'expression de la Distribution Normale Standard:

$$\mathcal{N}(z|0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Indice : Afin de trouver ces limites, vous pouvez utiliser l'intégration par parties.

Problème 3 (**Un exemple du Théorème Central Limite et des Queues de la Distribution Normale — 40 points**).

Reconsidérons l'exemple du Problème 1. Définissons S_N comme une somme de variables aléatoires indépendantes

$$S_N = \sum_{i=1}^{N} X_i \ ,$$

où les X_i sont des variables aléatoires indépendantes de Bernouilli avec paramètre $p = 1/2$, c'est-à-dire $\Pr\{X_i = 0\} = \Pr\{X_i = 1\} = 1/2$. Noter que ces X_i suivent la fonction indicatrice des faces. Utiliser le théorème limite central et les bornes obtenues au Problème 2 pour montrer que nous pouvons nous attendre à une probabilité d'avoir au moins $\frac{3}{4}N$ faces plus petite que

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{N}{8}}.$$

Note : Dans ce problème, notre objectif est de montrer qu'en utilisant le théorème central limite, nous pouvons obtenir un taux de décroissance plus rapide (dans le problème ci-dessus, *exponentiellement* rapide) que la borne de Bienaymé-Tchebychev trouvée dans le Problème 1 (qui nous donne un taux de décroissance *linéaire*).